

# Notice of Allowability

Application No.

10/731,164

Examiner

Angela M. Lie

Applicant(s)

KAYAHARA ET AL.

Art Unit

2163

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☒ This communication is responsive to 4/19/2007.
2. ☒ The allowed claim(s) is/are 2-5,7,8,10 and 11.
3. ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
  - a) ☒ All b) ☐ Some\* c) ☐ None of the:
    1. ☒ Certified copies of the priority documents have been received.
    2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
    3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

\* Certified copies not received: \_\_\_\_\_.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.

**THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.**

4. ☐ A SUBSTITUTE OATH OR DECLARATION must be submitted. Note the attached EXAMINER'S AMENDMENT or NOTICE OF INFORMAL PATENT APPLICATION (PTO-152) which gives reason(s) why the oath or declaration is deficient.
  5. ☐ CORRECTED DRAWINGS ( as "replacement sheets") must be submitted.
    - (a) ☐ including changes required by the Notice of Draftsperson's Patent Drawing Review ( PTO-948) attached
      - 1) ☐ hereto or 2) ☐ to Paper No./Mail Date \_\_\_\_\_.
    - (b) ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date \_\_\_\_\_.
- Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).
6. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

## Attachment(s)

1. ☒ Notice of References Cited (PTO-892)
2. ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3. ☐ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date \_\_\_\_\_
4. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material

5. ☐ Notice of Informal Patent Application
6. ☐ Interview Summary (PTO-413), Paper No./Mail Date \_\_\_\_\_
7. ☒ Examiner's Amendment/Comment
8. ☐ Examiner's Statement of Reasons for Allowance
9. ☐ Other \_\_\_\_\_



### EXAMINER'S AMENDMENT

1. An examiner's amendment to the record appears below. Should the changes and/or additions be unacceptable to applicant, an amendment may be filed as provided by 37 CFR 1.312. To ensure consideration of such an amendment, it MUST be submitted no later than the payment of the issue fee.
2. Authorization for this examiner's amendment was given in a telephone interview with Bogdan Zinchenko on June 8, 2007.
3. Please amend claims 2, 7, 8, 10 and 11 as shown in the attached amendment faxed in by the applicant on June 8<sup>th</sup>, 2007.
4. The Examiner's Amendment has been made in order to clarify minor informalities and overcome possible 35 U.S.C 112 second paragraph rejection, and therefore place this application in condition for allowance.

### *Allowable Subject Matter*

5. **Claims 2-5, 7, 8, 10 and 11 are allowed.**
6. The following is an examiner's statement of reasons for allowance:  
**As to claims 2, 7, 8, 10 and 11**, the prior art fails to teach a document extracting apparatus and method comprising the steps of: acquiring a plurality of documents from an information source, according to a user-specific criteria, computing all degrees of similarity between the plurality of documents, and express the degrees of similarity in a

Art Unit: 2163

symmetric matrix, computing all combinations of the degrees computed between plurality of documents, computing a sum of degrees of similarity between all of the documents and extracting documents constituting the combination with the smallest sum of the degrees of similarity among the plurality of documents constituting the respective combinations, wherein the similarity between the documents is computed in the manner as disclosed in claims 2, 7, 8, 10 and 11.

**As to claims 3-5**, those claims are allowed by the virtue of their dependency upon allowed claims.

7. Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

#### **The Prior Art**

8. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

- **Gomes et al (US Patent 6615209)** disclose a method for comparing two documents for the similarity.
- **Seki et al (US Publication 20020143737)** disclose an information retrieval device capable of comparing two documents and detecting repetitions.

**Inquiry**

9. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Angela M. Lie whose telephone number is 571-272-8445. The examiner can normally be reached on M-F.

10. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Don Wong can be reached on 571-272-1834. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

11. Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.



**Angela M Lie**



**DON WONG  
SUPERVISORY PATENT EXAMINER  
TECHNOLOGY CENTER 2100**

Attachment regarding Examiner's Amendment

**OLIFF & BERRIDGE, PLC**

ATTORNEYS AT LAW

277 SOUTH WASHINGTON STREET  
ALEXANDRIA, VIRGINIA 22314

TELEPHONE: (703) 836-6400

FACSIMILE: (703) 836-2787

E-MAIL: EMAIL@OLIFF.COM

WWW.OLIFF.COM

June 8, 2007

**U.S. PATENT AND TRADEMARK OFFICE  
FACSIMILE TRANSMISSION COVER SHEET**

<b>To:</b> In re the Application of  Naoki KAYAHARA et al.  <b>Application No.:</b> 10/731,164  <b>Filed:</b> December 10, 2003  <b>For:</b> DOCUMENT EXTRACTING DEVICE, DOCUMENT EXTRACTING PROGRAM, AND DOCUMENT EXTRACTING METHOD	<b>AUTHORIZATION TO ENTER EXAMINER'S AMENDMENT</b>  <b>Group Art Unit:</b> 2163  <b>Docket No.:</b> 117915
--	--

**Examiner:** A. LIE**Facsimile:** (571) 273-8445**From:** BOGDAN A. ZINCHENKO**Prepared By:** BAZ**Number of Pages Sent (Including cover sheet):** 8**Comments:**

Applicants hereby authorize Examiner Lie to enter an Examiner's Amendment amending the claims as shown in the attached Proposed Examiner's Amendment.

**Sent by:** BAZ

This facsimile is intended only for the use of the U.S. Patent and Trademark Office and contains confidential information. If you are not the intended recipient, or the employee or agent responsible to deliver it to the intended recipient, you are notified that any review, dissemination, distribution or copying of this facsimile is prohibited. If you have received this facsimile in error, please immediately notify us by facsimile or telephone, and return the facsimile to us by mail at the above address.

Application Serial No. 10/731,164

**Proposed Examiner's Amendment**

1. (Canceled)

2. (Currently Amended) A document extracting apparatus, comprising:

a document acquiring device to acquire a plurality of documents from an information source, according to a user-specific criteria, to be candidates for extraction;

a similarity computing device to compute all degrees of similarity between the plurality of documents, and express the degrees of similarity in a symmetric matrix,

the similarity computing device comprising:

a character-string-dividing functional unit to divide each of the plurality of documents into predetermined character strings;

a character-string frequency computing functional unit to compute document vectors of the plurality of documents on the basis of a frequency of appearance of the predetermined character strings divided by the character-string-dividing functional unit; and

a mutual similarity computing functional unit to compute the degrees of similarity between the plurality of documents on the basis of the document vectors obtained from the character-string frequency computing functional unit;

a combination computing device to compute all combinations of ~~any number of documents from the degrees of similarity computed between~~ the plurality of documents;

a sum of degrees of similarity computing device to compute, with respect to all of the combinations, a sum of the degrees of similarity between all of the documents that constitute each combination, based on all of the degrees of similarity expressed in the symmetric matrix; and

## Application No. 10/731,164

a document extracting device to extract documents constituting the combination with the smallest sum of the degrees of similarity among the plurality of documents constituting the respective combinations.

3. (Previously Presented) The document extracting apparatus according to Claim 2, the character-string-dividing functional unit dividing each of the plurality of documents into predetermined character strings using any of the following character string division methods: a morphological analysis method, an n-gram method, and a stop-word method.

4. (Previously Presented) The document extracting apparatus according to Claim 2, the character-string frequency computing functional unit generating document vectors obtained by weighting each of the plurality of documents by a term frequency and inverse document frequency (TFIDF) weighting method on the basis of a frequency of appearance of the divided character strings.

5. (Previously Presented) The document extracting apparatus according to Claim 2, the mutual similarity computing functional unit computing degrees of similarity between the plurality of documents by a vector space method on the basis of the document vectors of the plurality of documents.

6. (Canceled)

7. (Currently Amended) A computer-readable media having a document extracting program allowing a computer to serve as:

a document acquiring device to acquire a plurality of documents from an information source, according to a user-specific criteria, to be candidates for extraction;

a similarity computing device to compute all degrees of similarity between the plurality of documents, and express the degrees of similarity in a symmetric matrix,

Application No. 10/731,164

the similarity computing device comprising:

a character-string-dividing function to divide each of the plurality of documents into predetermined character strings;

a character-string frequency computing function to compute document vectors of the plurality of documents on the basis of a frequency of appearance of the predetermined character strings divided by the character-string-dividing function; and

a mutual similarity computing function to compute the degrees of similarity between the plurality of documents on the basis of the document vectors obtained by the character-string frequency computing function;

a combination computing device to compute all combinations of ~~any number of documents from the degrees of similarity computed between~~ the plurality of documents;

a sum of degrees of similarity computing device to compute, with respect to all of the combinations, a sum of the degrees of similarity between all of the documents that constitute each combination, based on all of the degrees of similarity expressed in the symmetric matrix; and

a document extracting device to extract documents constituting the combination with the smallest sum of the degrees of similarity among the plurality of documents constituting the respective combinations.

8. (Currently Amended) A computer-readable media having a document extracting program allowing a computer to serve as:

a document acquiring device to acquire a plurality of documents from an information source, according to a user-specific criteria, to be candidates for extraction;



Application No. 10/731,164

a similarity computing device to compute all degrees of similarity between the plurality of documents, and express the degrees of similarity in a symmetric matrix,

the similarity computing device comprising:

a character-string-dividing function to divide each of the plurality of documents into character strings using any one of character string division methods;

a character-string frequency computing function to generate document vectors obtained by weighting each of the documents by a term frequency and inverse document frequency (TFIDF) weighting method on the basis of a frequency of appearance of the divided character strings; and

a mutual similarity computing function to compute the degrees of similarity between the plurality of documents by a vector space method on the basis of the document vectors of the plurality of documents,

a combination computing device to compute all combinations of ~~any number of documents from the~~ degrees of similarity computed between the plurality of documents;

a sum of degrees of similarity computing device to compute, with respect to all of the combinations, a sum of the degrees of similarity between all of the documents that constitute each combination, based on all of the degrees of similarity expressed in the symmetric matrix; and

a document extracting device to extract documents constituting the combination with the smallest sum of the degrees of similarity among the plurality of documents constituting the respective combinations.

9. (Canceled)

Application No. 10/731,164

10. (Currently Amended) A document extracting method, comprising:

acquiring a plurality of documents from an information source, according to a user-specific criteria, to be candidates for extraction;

dividing each of the documents into predetermined character strings, computing a frequency of appearance of the divided character strings, computing document vectors of the plurality of documents on the basis of the frequency of appearance of the predetermined character strings, and computing the degrees of similarity between the plurality of documents using the document vectors;

~~computing all degrees of similarity between the plurality of documents, and~~

~~expressing the degrees of similarity in a symmetric matrix;~~

~~computing all combinations of any number of documents from the degrees of similarity computed between the plurality of documents;~~

computing, with respect to all of the combinations, a sum of the degrees of similarity between all of the documents that constitute each combination, based on all of the degrees of similarity expressed in the symmetric matrix; and

extracting documents constituting the combination with the smallest sum of the degrees of similarity among the plurality of documents constituting the respective combinations;

~~dividing each of the documents into predetermined character strings, computing a frequency of appearance of the divided character strings, computing document vectors of the plurality of documents on the basis of the frequency of appearance of the predetermined character strings, and then computing the degrees of similarity between the plurality of documents using the document vectors.~~

Application No. 10/731,164

11. (Currently Amended) A document extracting method, comprising:

acquiring a plurality of documents from an information source, according to a user-specific criteria, to be candidates for extraction;

dividing each of the plurality of documents into predetermined character strings using any one of character string division methods, including a morphological analysis method, an n-gram method, and a stop-word method, computing document vectors of the plurality of documents by weighting each of the documents by a term frequency and inverse document frequency (TFIDF) weighting method on the basis of a frequency of appearance of the divided predetermined character strings, and computing the degrees of similarity between the plurality of documents using a vector space method on the basis of the document vectors;

computing all degrees of similarity between the plurality of documents, and expressing the degrees of similarity in a symmetric matrix;

computing all combinations of any number of documents from the degrees of similarity computed between the plurality of documents;

computing, with respect to all of the combinations, a sum of the degrees of similarity between all of the documents that constitute each combination, based on all of the degrees of similarity expressed in the symmetric matrix; and

extracting documents constituting the combination with the smallest sum of the degrees of similarity among the plurality of documents constituting the respective combinations.

~~dividing each of the plurality of documents into predetermined character strings using any one of character string division methods, including a morphological analysis method, an n-gram method, and a stop-word method, computing document vectors of the plurality of~~

Application No. 10/731,164

~~documents by weighting each of the documents by a term frequency and inverse document frequency (TFIDF) weighting method on the basis of a frequency of appearance of the divided predetermined character strings, and computing the degrees of similarity between the plurality of documents using a vector space method on the basis of the document vectors.~~

12-14. (Canceled)